

Combining Ground and Satellite Data

Antoine Zelenka

Swiss Meteorological Institute, 8044 Zurich, Switzerland

External data inputs, such as satellite imagery, can greatly improve spatial interpolation between the nodes of radiation measuring networks. Reciprocally, network measurements can improve the accuracy of satellite-based radiation estimates. An adequate technique for these data fusion processes is the bi-variate geostatistical interpolation called co-kriging.

Consider the satellite estimates, g , as the better sampled, but less accurate co-variable capable of refining the spatial resolution of an interpolation procedure based on the less well sampled, but more accurate principal variable, G , delivered by the radiometric network. Global irradiation G at any location x is then obtained from the bi-linear estimator

$$G(x) = \sum w(x_i, x) G(x_i) + \sum v(x_k, x) g(x_k) \quad (1)$$

where the first sum stands for the contribution of the network, with $G(x_i)$ being the measured value at the site x_i , while the second accounts for the contribution of the satellite-based estimates g at each pixel x_k . The first sum runs over those network sites i , the second over those pixels k , which lie within the pre-chosen search domain around x . Unbiasedness is enforced by requiring that $\sum w_i = 1$ and

$\sum v_k = 0$. Both conditions, together with that of minimum variance of estimation, allow a unique determination of the weights w_i and v_k under only very general assumptions.

The weights v predominate in the regions where g and G correlate well, while the weights w predominate when the latter correlation degrades. Thus, the resulting $G(x)$ matches the network's values (within prescribed limits) but respects as much as possible the fine structure defined by the satellite.

The main advantage of co-kriging is that it yields a value for the variance of estimation for every x , a value essential to any solar system planning and design activity. Its main drawback is that it cannot be run completely automatically, because it requires expert selection of functions describing the structure of the fields $G(x_i)$ and $g(x_k)$.

Continuous, stand-alone estimations, as in the case of hourly irradiation, have therefore to benefit from ground observations in some other way. For example, if the network measures also the direct beam at several sites, then essential turbidity information can be input on-line to the satellite model.

COMBINING GROUND AND SATELLITE DATA

Antoine Zelenka

USING GROUND MEASUREMENTS TO IMPROVE SATELLITE ESTIMATES

NETWORK MEASUREMENTS

Radiation monitoring networks (or, at worst: meteorol. networks delivering insolation-related quantities)

Consider simple, robust models with only one to two parameters, delivering estimates for hourly irradiation with 20%-25% RMSE* as compared to the network measurements.

* „state of the art“ values (Gautier and Landsfeld obtain 18%. Gautier C. and Landsfeld M., 1997. Surface solar radiation flux and cloud radiative forcing for the ARM/SGP: A satellite, surface observations and radiative transfer model study. *J. Atmos. Sci.* 54, 1289-1307.)

IMPROVING THE ESTIMATES:

Require that the satellite estimates match the ground measurements at the network nodes within certain limits.

Note: This is not equivalent to shifting the satellite-derived $G(x)$ surface rigidly „up“ or „down“ above the x plane in order to remove the estimation bias (leaves the RMSE almost unchanged if the bias is not too large). Rather, add local trends to the surface, so that it matches the nodes without loosing its fine structure.

POINT MEASUREMENTS

Elaborate programmes at carefully selected sites during special observing periods:

Surface SW & LW irradiance and components (\Rightarrow turbidity), radiosondes, LIDARs, SODARs, RADARs, aircrafts, etc. provide thorough characterisation of the vertical structure of the atmosphere, as well as micro- and macrophysical properties of clouds.

\Rightarrow integration of equation of radiative transfer along the line of sight to the satellite and for the downward flux density at the surface.

Charlock T. P. and Alberta T. L., 1996. The CERES/ARM/ GEWEX Experiment (CAGEX) for the retrieval of radiative fluxes with satellite data. *Bull. Amer. Meteor. Soc.*, 77, 2673-2683.

Chou M.-D. and Zhao W., 1997. Estimation and model validation of surface solar radiation and cloud radiative forcing using TOGA COADS measurements. *J. Climate*, 10, 610-620.

Inadequate for large amount of data as regional insolation mapping with high spatial resolution or for long-term high temporal resolution datasets.

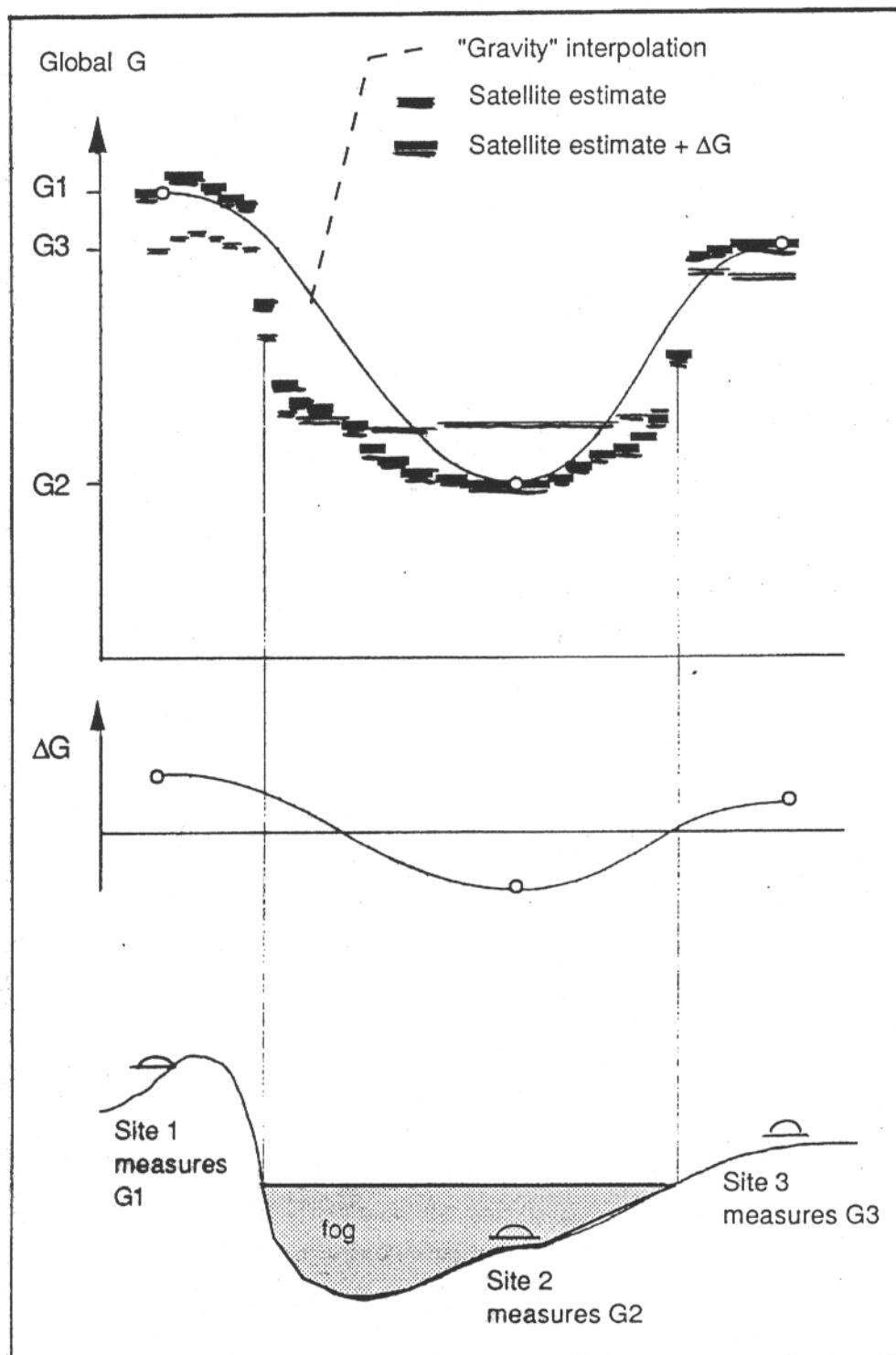


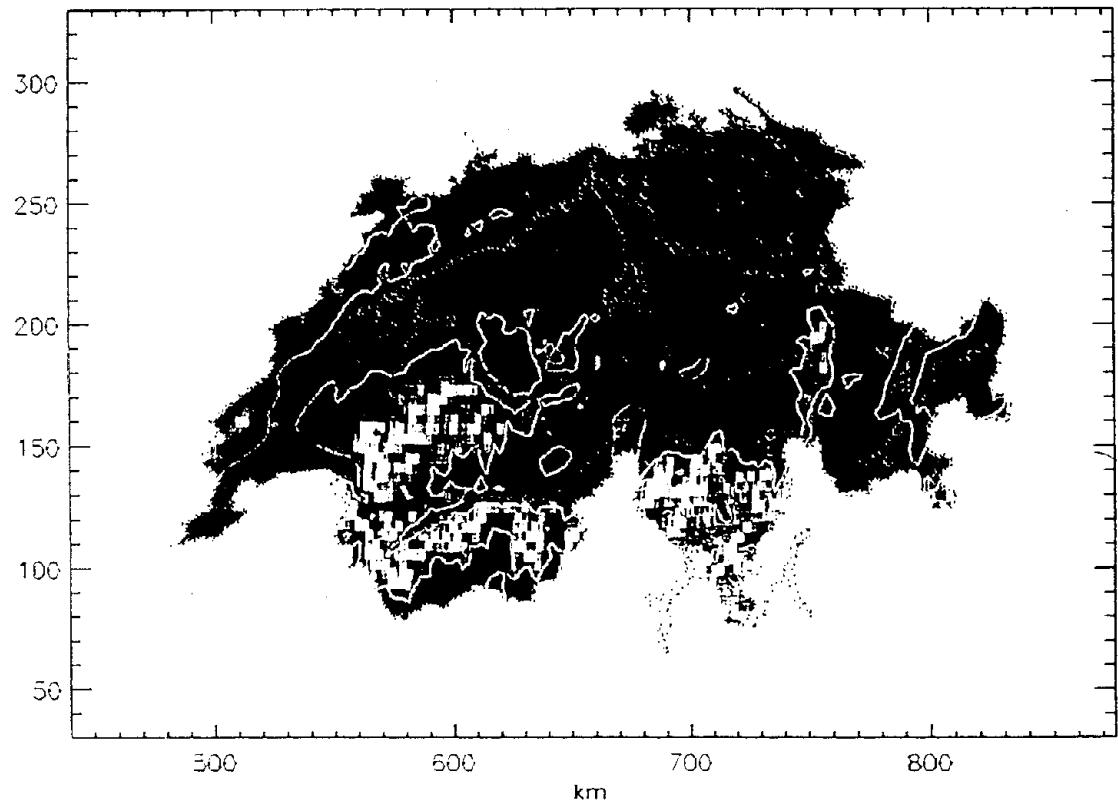
Fig. 5.3.1: Mere interpolation between sites 1, 2 and 3 misses the sharp transitions at the fog boundaries. Satellite estimates render them correctly in location but not necessarily in absolute value. Application of an interpolated, i.e., location-dependent correction $\Delta G = G(\text{meas}) - G(\text{sat})$ leads to predictions which reproduce the measurements without real loss of spatial resolution. The discrete, discontinuous shape of the satellite estimates is intended to mimic the pixel size of the satellite image.

VPP

azk

E13

Filename :	tmp.ps
Device :	CLC800.PTS
Form :	A4
Code :	ISOLATIN1
Print format :	CRLF
Job length :	1'437'759 bytes
Job got from	129.132.1.24
Job created by	UNIX,APRIL 90, 129.132.1.24
Job queued at	30-JAN-1997 12:31:58
Job printed at	30-JAN-1997 12:31:59



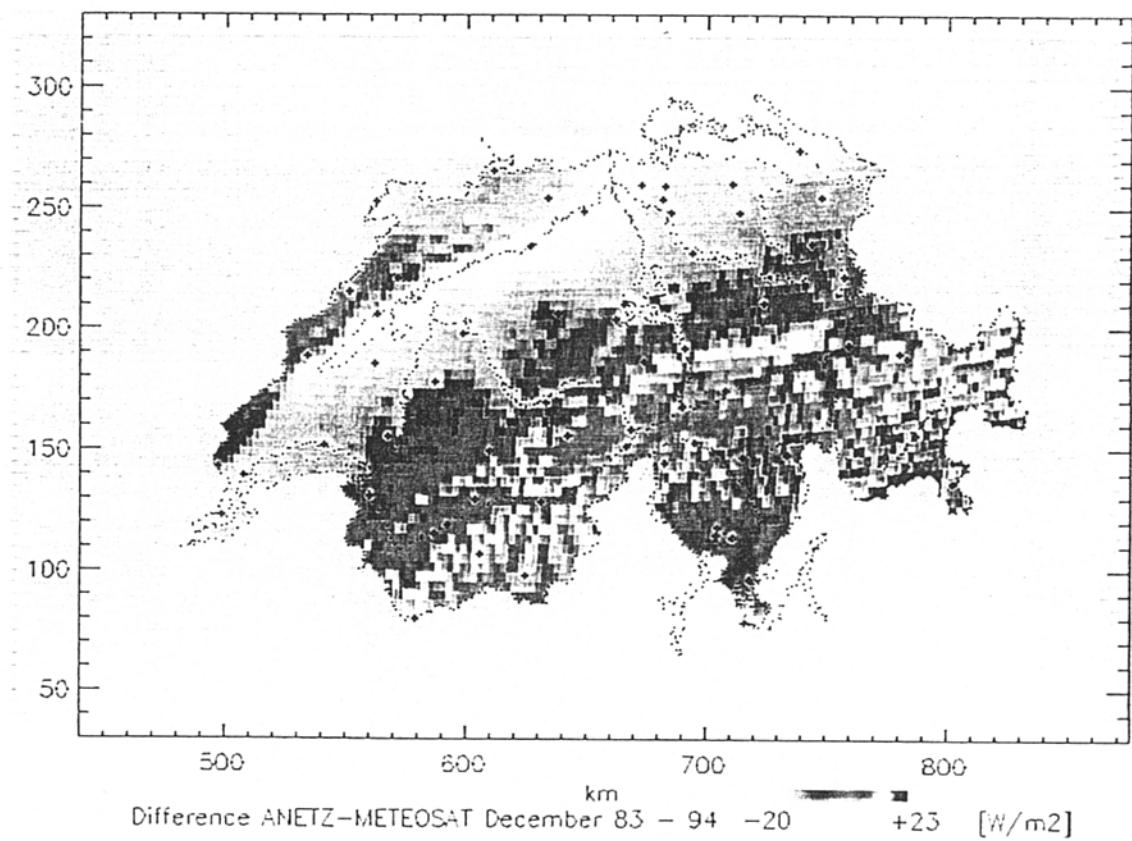
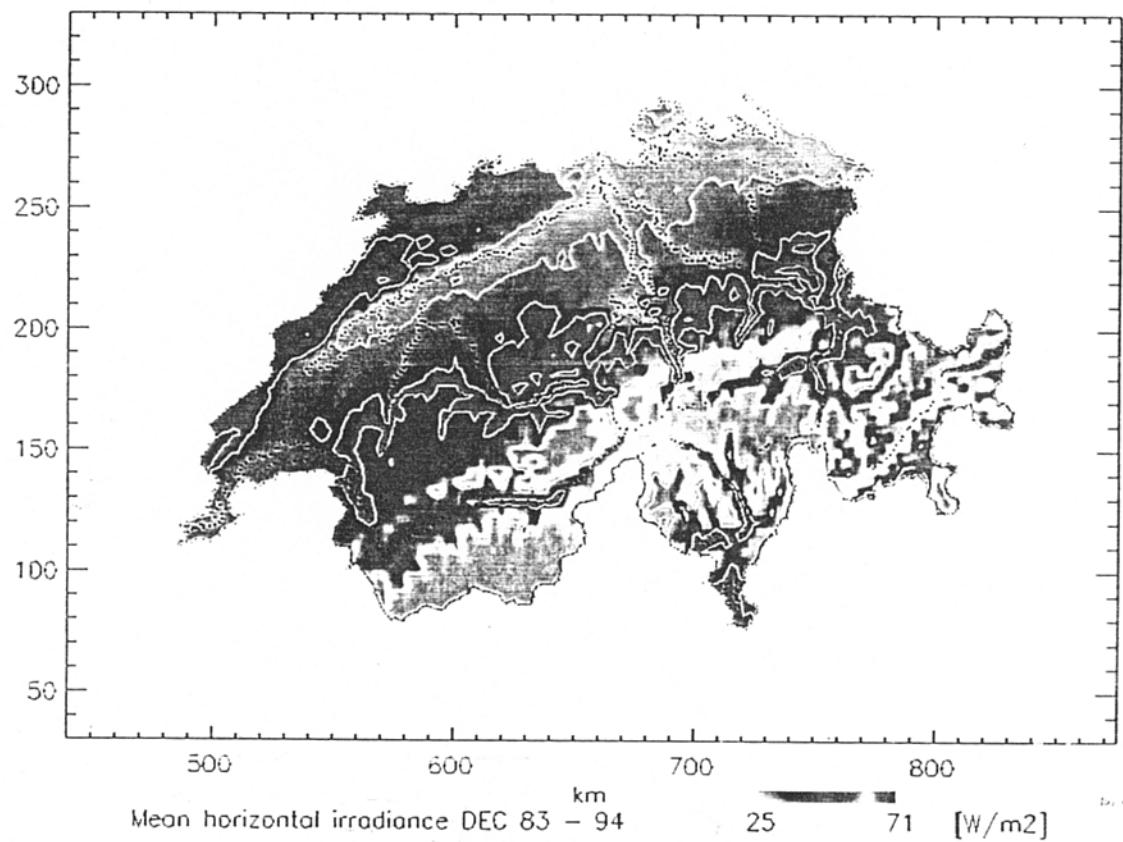
azk

E13

Filename : tmp.ps
Device : CLC800.PTS
Form : A4
Code : ISOLATIN1
Print format : CRLF
Job length : 1'437'759 bytes
Job got from 129.132.1.24
Job created by UNIX,APRIL 90, 129.132.1.24
Job queued at 8-APR-1997 09:45:54
Job printed at 8-APR-1997 09:45:56

VPP ist ein Produkt und eine Dienstleistung der Informatikdienste der ETH Zürich

ETZSPEZ CLC800 Job-02



VPP

azk

E13

Filename : tmp.ps
Device : CLC500.PTS
Form : A4
Code : ISOLATIN1
Print format : CRLF
Job length : 1'437'758 bytes
Job queued at 27-SEP-1995 11:34:58
Job started at 27-SEP-1995 19:49:55

Kriging in the Intrinsic Case

Hypotheses: Field $Z(x)$ is "weakly stationary of order two"

Introducing the Variogram function $\gamma(h)$

$$E[Z(x+h) - Z(x)] = 0$$

$$\text{Var}[Z(x+h) - Z(x)] = 2\gamma(h) \quad (\text{independent of } x)$$

Formulation:

Weights Measured values at network nodes

Linear Estimator $Z_0^* = \sum_{\alpha} \lambda^{\alpha} Z(x_{\alpha})$

No bias $E[Z_0^* - Z_0] = 0$

Optimum $\text{Var}[Z_0^* - Z_0] \text{ minimal}$

Ordinary Kriging equations: Build empirical variogram $\gamma(h)$

$$\sum_{\beta} \lambda^{\beta} \gamma_{\alpha\beta} + \mu_0 = \gamma_{\alpha 0}$$

$$\sum_{\alpha} \lambda^{\alpha} = 1$$

$$(\sigma^2)^* = -\gamma_{00} + \sum_{\alpha} \lambda^{\alpha} \gamma_{\alpha 0} + \mu_0$$

Variance of
Estimation Z_0^*

Limits within which Co-K results have to honour measured values at the network nodes:

Recall the results of the *effective RMSE* presentation: the limits are defined by the intercomparison noise originating from

- the time/space disparity between the satellite-derived and the ground data,
- from measuring errors (ground and satellite),
- from the genuine micro-variability of the irradiance field at the sub-pixel scale.

This **intercomparison noise** has been found to amount to **about 19% RMSE**. The result of the data fusion process can deviate by this amount from the network's measurements.

Co-K drawback: Variography requires expert knowledge and should not be automated.

=> Co-K is not suited for unattended permanent hourly retrievals.

Acceptable alternative:

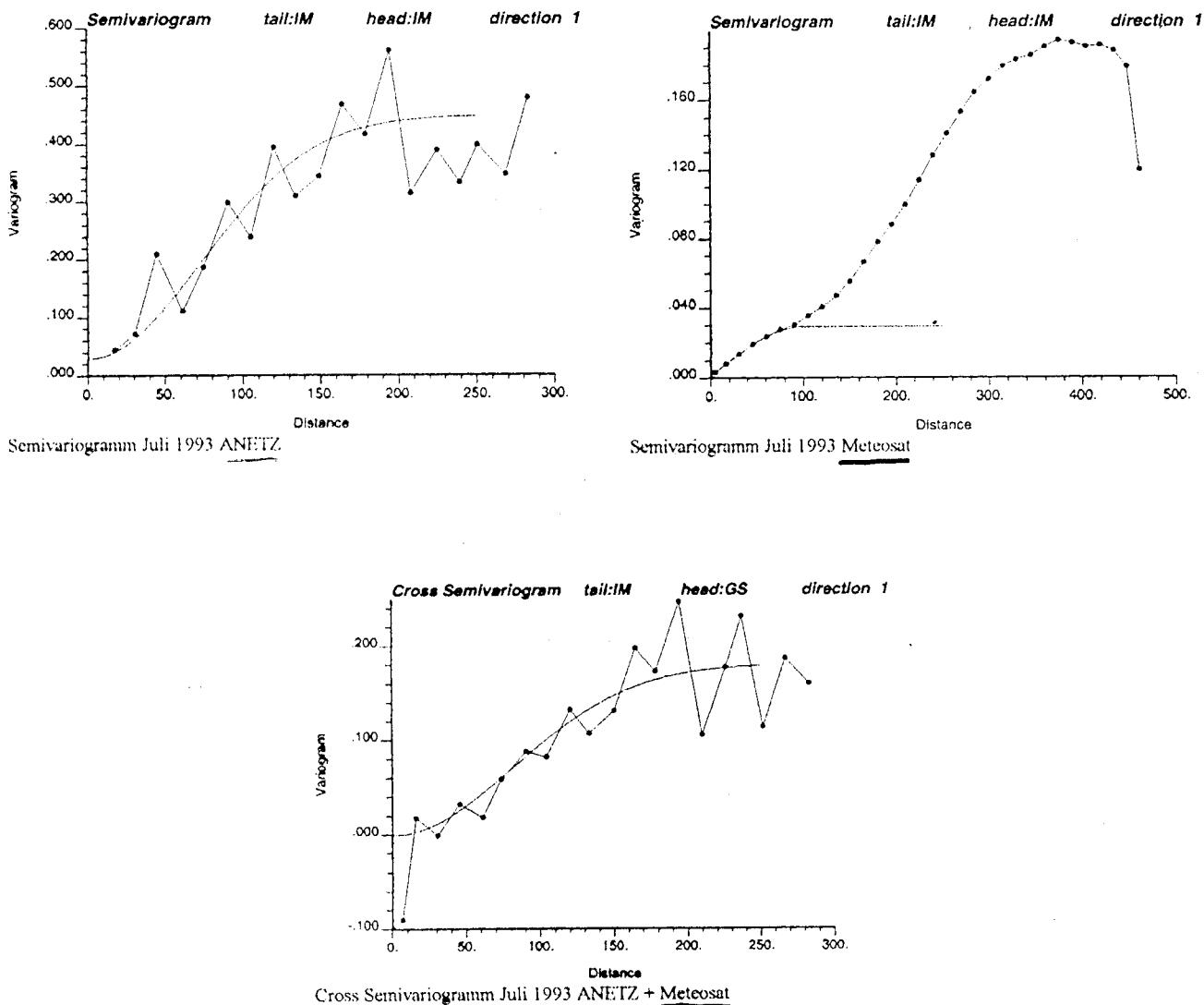
Use pre-chosen weights to interpolate the differences ground minus satellite at the network nodes. Add map of differences to satellite map (as in example above).

6.2. Cokriging für den Juli 1993

Für die Interpolation mit Cokriging wurde das Programm *cokb3dm.exe* aus GSLIB benutzt. Um die Interpolation durchzuführen, müssen die beiden Variogramme der beiden Datensätze sowie deren Crossvariogramm vorhanden sein. Diese drei Variogramme müssen die Cauchi-Schwarz-Relation

$$|\gamma_{12}(h)| \leq \sqrt{[\gamma_1(h)\gamma_2(h)]}$$

für jeden Punkt des Variogrammes erfüllen, was hier vollumfänglich der Fall ist.

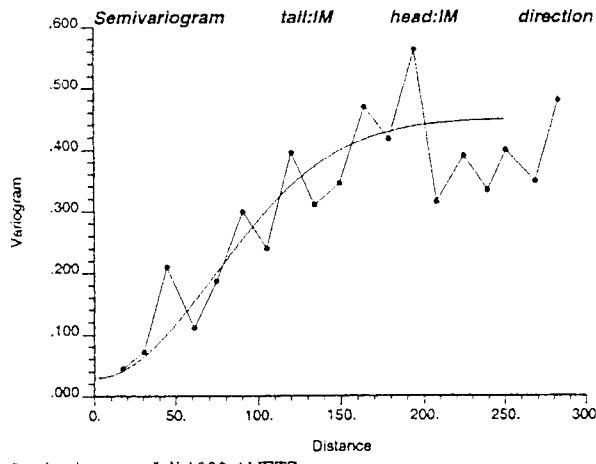


6.2. Cokriging für den Juli 1993

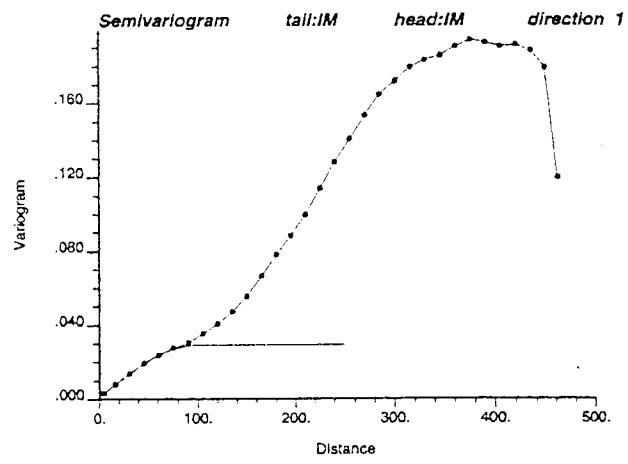
Für die Interpolation mit Cokriging wurde das Programm *cokb3dm.exe* aus GSLIB benutzt. Um die Interpolation durchzuführen, müssen die beiden Variogramme der beiden Datensätze sowie deren Crossvariogramm vorhanden sein. Diese drei Variogramme müssen die Cauchi-Schwarz-Relation

$$|\gamma_{12}(h)| \leq \sqrt{[\gamma_1(h)\gamma_2(h)]}$$

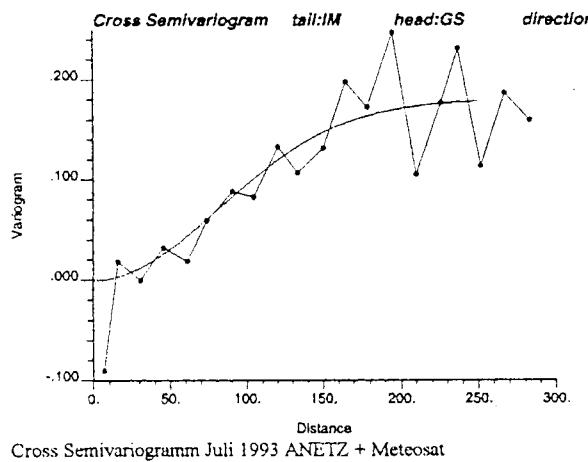
für jeden Punkt des Variogrammes erfüllen, was hier vollumfänglich der Fall ist.



Semivariogramm Juli 1993 ANETZ



Semivariogramm Juli 1993 Meteosat



Cross Semivariogramm Juli 1993 ANETZ + Meteosat

Die Interpolation brachte die beiden Karten auf den nächsten zwei Seiten hervor. Die erste ist die eigentliche Interpolation in kWh/m^2 , sie zeigt einige Änderungen im Vergleich zur Kriging-Karte, speziell im kleinräumigen Bereich sind feinere Strukturen vorhanden. Im Grossen und Ganzen sind sich die beiden Karten aber sehr ähnlich. Die zweite Karte stellt die dazugehörige Schätzvarianz dar, die zur Fehlerabschätzung benutzt wird. Im Vergleich zum Kriging sind die Werte durchwegs kleiner, was auf grösitere Genauigkeit im Schätzprozess hindeutet. Allerdings sind die Unterschiede geringer als erwartet. Speziell aber die hohe Schätzvarianz der beiden Dreiecke Basel-Chasseral-Bern sowie Wädenswil-Chur-Disentis hat stark abgenommen, was zeigt dass die Meteosatdaten dort einen nicht zu unterschätzenden Einfluss auf die Schätzung haben. Somit lässt sich die Aussage machen, dass eine

Co-kriging

Bi-variate geostatistical interpolation with bi-linear estimator

$$G(\underline{x}) = \sum w(\underline{x}_i, \underline{x}) G(\underline{x}_i) + \sum v(\underline{x}_k, \underline{x}) g(\underline{x}_k)$$

First sum accounts for the contribution of the network measurements G at stations \underline{x}_i

Second sum accounts for the contribution of the satellite-based estimates g at each pixel \underline{x}_k .

The first sum runs over those network sites i , the second over those pixels k , which lie within a pre-chosen search domain around \underline{x} .

Unbiasedness is enforced by requiring that $\sum w_i = 1$ and $\sum v_k = 0$. Both conditions, together with that of minimum variance of estimation, allow a *unique* determination of the weights w_i and v_k under only very general assumptions.

The weights v predominate in the regions where g and G correlate well, while the weights w predominate when the latter correlation degrades.

When main and co-variable correlate perfectly, then co-kriging reduces to kriging interpolation of the differences between both variables.

The main advantage of co-kriging is that it yields a value for ***the variance of estimation*** for every \underline{x} , a value essential to any solar system planning and design activity

Promising alternative

Consider the following, extremely simple satellite model:

$$\tau_{\text{atm}}(n) = \tau_{\text{atm}}(0) (1 - n)$$

where $\tau_{\text{atm}}(n)$ and $\tau_{\text{atm}}(0)$ are the atmospheric transmittances for the cloudy and clear sky, respectively. The cloudiness is expressed with the cloud index n (Cano *et al.*, 1986)

$$n = (\alpha_{\text{toa}} - \alpha_{\text{toa,min}}) / (\alpha_{\text{toa,max}} - \alpha_{\text{toa,min}})$$

α_{toa} is the instantaneous planetary albedo,

$\alpha_{\text{toa,min}}$ corresponds to a clear, clean and dry sky,

$\alpha_{\text{toa,max}}$ corresponds to a heavily overcast sky.

Clear sky transmissivity

Model of Kasten *et al.* (1984), derived from 10 years of observations by the radiometric network of the German Weather Service:

$$\tau_{\text{atm}}(0) = 0.84 \exp(-0.027 T_L m)$$

Suggestion:

Use an RSR network to derive Linke's turbidity factor T_L

Promising alternative

Consider the following, extremely simple satellite model:

$$\tau_{\text{atm}}(n) = \tau_{\text{atm}}(0) (1 - n)$$

where $\tau_{\text{atm}}(n)$ and $\tau_{\text{atm}}(0)$ are the atmospheric transmittances for the cloudy and clear sky, respectively. The cloudiness is expressed with the cloud index n (Cano *et al.*, 1986)

$$n = (\alpha_{\text{toa}} - \alpha_{\text{toa,min}}) / (\alpha_{\text{toa,max}} - \alpha_{\text{toa,min}})$$

α_{toa} is the instantaneous planetary albedo,

$\alpha_{\text{toa,min}}$ corresponds to a clear, clean and dry sky,

$\alpha_{\text{toa,max}}$ corresponds to a heavily overcast sky.

Clear sky transmissivity

Model of Kasten *et al.* (1984), derived from 10 years of observations by the radiometric network of the German Weather Service:

$$\tau_{\text{atm}}(0) = 0.84 \exp(-0.027 T_L m)$$

Suggestion:

Use an RSR network to derive Linke's turbidity factor T_L

PERFORMANCE EVALUATION

Swiss network 4/94

Daily Data

NETWORK KRIGING	—	17%
SATELLITE ALONE	—	17%
CO-KRIGING	—	<u>13.5 %</u>